

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-249957

(43)Date of publication of application : 05.09.2003

(51)Int.Cl.

H04L 12/56
G10L 19/00
G10L 19/12
H04L 1/00

(21)Application number : 2002-045839

(71)Applicant : NIPPON TELEGR & TELEPH CORP <NTT>

(22)Date of filing : 22.02.2002

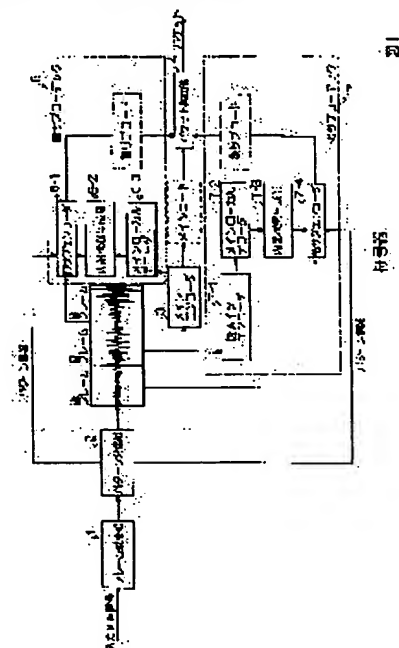
(72)Inventor : MORINAGA TORU
MANO KAZUNORI

(54) METHOD AND DEVICE FOR CONSTITUTING PACKET, PROGRAM FOR CONSTITUTING PACKET, AND METHOD AND DEVICE FOR PACKET DISASSEMBLY, PROGRAM FOR PACKET DISASSEMBLY

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a method and a device for constituting packets for reducing degradation of quality caused by the loss of packets by adding a small amount of information.

SOLUTION: A pattern classifying part 2 composes a compensation signal of a preceding (succeeding) frame by extrapolating repeat code of a sound signal of a current frame or the feature quantity of this code and compares the compensation signal with the preceding (succeeding) frame signal to discriminate whether a preceding (succeeding) frame artificial signal can be generated from the current frame or not. When it is discriminated that it cannot be generated, a preceding (succeeding) sub-code is generated on the basis of the preceding (succeeding) frame signal by a preceding (succeeding) sub-encoder 6-1 (7-4), and a packet constituting part 4 adds the preceding (succeeding) sub-code to the main code of the current frame encoded by a main encoder 3.



LEGAL STATUS

[Date of request for examination]

10.02.2004

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-249957

(P2003-249957A)

(43) 公開日 平成15年9月5日 (2003.9.5)

(51) Int.Cl. ⁷	識別記号	F I	ターマコード* (参考)
H 0 4 L 12/56	3 0 0	H 0 4 L 12/56	3 0 0 Z 5 D 0 4 5
	2 3 0		2 3 0 Z 5 K 0 1 4
G 1 0 L 19/00		1/00	B 5 K 0 3 0
19/12		G 1 0 L 9/14	S
H 0 4 L 1/00			N
審査請求 未請求 請求項の数 9 O L (全 11 頁)			

(21) 出願番号 特願2002-45839 (P2002-45839)

(22) 出願日 平成14年2月22日 (2002.2.22)

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区大手町二丁目3番1号

(72) 発明者 森永 徹

東京都千代田区大手町二丁目3番1号 日

本電信電話株式会社内

(72) 発明者 間野 一則

東京都千代田区大手町二丁目3番1号 日

本電信電話株式会社内

(74) 代理人 100066153

弁理士 草野 卓 (外1名)

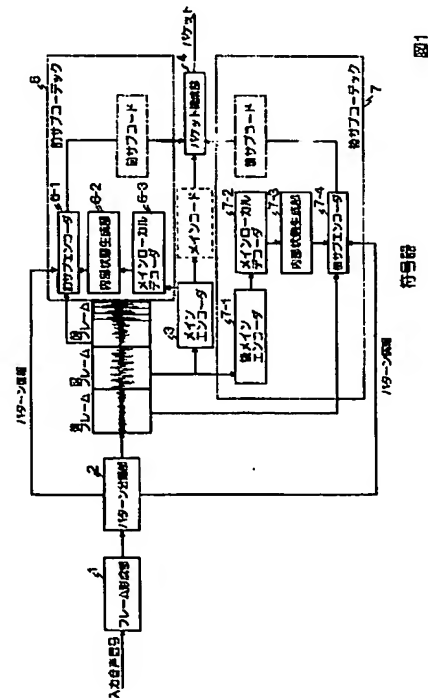
最終頁に続く

(54) 【発明の名称】 パケット構成方法及び装置、パケット構成プログラム、並びにパケット分解方法及び装置、パケット分解プログラム

(57) 【要約】

【課題】 少ない情報量の付加でパケット消失による品質劣化を低減するパケット構成方法、装置を提供する。

【解決手段】 パターン分類部2は現フレームの音声信号の繰り返し又は該符号の特徴量の外挿により前（後）フレームの補償信号を合成し、前（後）フレーム信号と比較することにより現フレームから前（後）フレーム疑似信号を作ることができるか否かを判断し、作ることができないと判断された場合には前（後）フレーム信号をもとに前（後）サブエンコーダ6-1（7-4）により前（後）サブコードを生成し、パケット構成部4はメインエンコーダ3で符号化した現フレームのメインコードに前（後）サブコードを付加する。



【特許請求の範囲】

【請求項 1】 音声信号をフレームごとに符号化した符号をパケットに格納するパケット構成方法において、
現フレームの音声信号の繰り返しまたは該符号の特徴量の外挿により前フレーム及び後フレームの補償信号を合成する過程と、

前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より大きく、後フレームの信号波形と前記後フレームの補償信号との歪が所定の閾値より大きい場合、現フレームと、前フレームと後フレームの符号を含めてパケットを構成する過程と、

前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より大きく、後フレームの補償信号との歪が所定の閾値より小さい場合、現フレームと前フレームの符号と前フレームを示す符号を含めてパケットを構成する過程と、

前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より小さく、後フレームの信号波形と前記後フレームの補償信号との歪が所定の閾値より大きい場合、現フレームと後フレームの符号と後フレームを示す符号を含めてパケットを構成する過程と、を有することを特徴とするパケット構成方法。

【請求項 2】 請求項 1 に記載のパケット構成方法において、

前記前フレーム又は後フレームの符号は、前記現フレームに対する符号化方法とは異なる符号化方法で前記前フレーム又は後フレームの音声信号に対して前記現フレームの符号化による復号信号に基づく内部状態変数を用いて生成することを特徴とするパケット構成方法。

【請求項 3】 音声信号をフレームごとに符号化した符号をパケットに格納するパケット構成装置において、
現フレームの音声信号の繰り返しまたは該符号の特徴量の外挿により前フレーム及び後フレームの補償信号を合成する手段と、
前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より大きく、後フレームの信号波形と前記後フレームの補償信号との歪が所定の閾値より大きい場合、現フレームと、前フレームと後フレームの符号を含めてパケットを構成し、前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より大きく、後フレームの補償信号との歪が所定の閾値より小さい場合、現フレームと前フレームの符号と前フレームを示す符号を含めてパケットを構成し、前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より小さく、後フレームの信号波形と前記後フレームの補償信号との歪が所定の閾値より大きい場合、現フレームと後フレームの符号と後フレームを示す符号を含めてパケットを構成する手段と、を備えたことを特徴とするパケット構成装置。

【請求項 4】 請求項 3 に記載のパケット構成装置におい

て、

前記前フレーム又は後フレームの符号は、前記現フレームに対する符号化方法とは異なる符号化方法で前記前フレーム又は後フレームの音声信号に対して前記現フレームの符号化による復号信号に基づく内部状態変数を用いて生成することを特徴とするパケット構成装置。

【請求項 5】 音声信号をフレームごとに符号化した符号をパケットに格納する処理をコンピュータに実行させるパケット構成プログラムにおいて、

10 現フレームの音声信号の繰り返しまたは該符号の特徴量の外挿により前フレーム及び後フレームの補償信号を合成する処理と、

前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より大きく、後フレームの信号波形と前記後フレームの補償信号との歪が所定の閾値より大きい場合、現フレームと、前フレームと後フレームの符号を含めてパケットを構成する処理と、

20 前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より大きく、後フレームの補償信号との歪が所定の閾値より小さい場合、現フレームと前フレームの符号と前フレームを示す符号を含めてパケットを構成する処理と、

前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より小さく、後フレームの信号波形と前記後フレームの補償信号との歪が所定の閾値より大きい場合、現フレームと後フレームの符号と後フレームを示す符号を含めてパケットを構成する処理と、をコンピュータに実行させるパケット構成プログラム。

30 【請求項 6】 請求項 5 に記載のパケット構成プログラムにおいて、

前記前フレーム又は後フレームの符号は、前記現フレームに対する符号化方法とは異なる符号化方法で前記前フレーム又は後フレームの音声信号に対して前記現フレームの符号化による復号信号に基づく内部状態変数を用いて生成することを特徴とするパケット構成プログラム。

【請求項 7】 パケット毎に格納されたフレーム毎の符号を復号化して音声信号を再生するパケット分解方法において、

パケットが消失したか否かを判定する過程と、

40 現パケットが消失した場合、

前パケットが後フレーム符号を含むとき、当該後フレーム符号を復号して現パケットの音声信号を再生する過程と、

後パケットが前フレーム符号を含むとき、当該前フレーム符号を復号して現パケットの音声信号を再生する過程と、

50 前パケットも後パケットも前フレーム符号も後フレーム符号も含まず現フレーム符号を含むとき、前後いずれか一方のパケットの当該現フレーム符号の復号信号の繰り返し又は該信号の特徴量の補間もしくは外挿により現パ

3

ケットの音声信号を再生する過程と、を有することを特徴とするケット分解方法。

【請求項8】ケット毎に格納されたフレーム毎の符号を復号化して音声信号を再生するケット分解装置において、

ケットが消失したか否かを判定する手段と、

現ケットが消失した場合、

前ケットが後フレーム符号を含むとき、当該後フレーム符号を復号して現ケットの音声信号を再生し、後ケットが前フレーム符号を含むとき、当該前フレーム符号を復号して現ケットの音声信号を再生し、前ケットも後ケットも前フレーム符号も後フレーム符号も含まず現フレーム符号を含むとき、前後いずれか一方のケットの当該現フレーム符号の復号信号の繰り返し又は該信号の特徴量の補間もしくは外挿により現ケットの音声信号を再生する手段と、を備えたことを特徴とするケット分解装置。

【請求項9】ケット毎に格納されたフレーム毎の符号を復号化して音声信号を再生する処理をコンピュータに実行させるケット分解プログラムにおいて、

ケットが消失したか否かを判定する処理と、

現ケットが消失した場合、

前ケットが後フレーム符号を含むとき、当該後フレーム符号を復号して現ケットの音声信号を再生する処理と、

後ケットが前フレーム符号を含むとき、当該前フレーム符号を復号して現ケットの音声信号を再生する処理と、

前ケットも後ケットも前フレーム符号も後フレーム符号も含まず現フレーム符号を含むとき、前後いずれか一方のケットの当該現フレーム符号の復号信号の繰り返し又は該信号の特徴量の補間もしくは外挿により現ケットの音声信号を再生する処理と、をコンピュータに実行させるケット分解プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は音声信号を圧縮符号化してケットに収容する方法及び装置、もしくは伝送されたケットに収容された符号から音声信号を復号する方法及び装置に関する。

【0002】

【従来の技術】移動体通信やVoIP(Voice over IP)に代表されるように、ケット通信によって音声とデータを統合的に扱う事が可能となる。ケット音声通信における問題点として、符号化による音声の劣化、遅延、ケット消失があげられる。通信路が広帯域化、高速化されることにより、符号化による劣化、遅延は解消されるが、ケット消失は通信容量が増えても生じる可能性がありうる問題である。ケット消失が起こる原因として次のものがあげられる。まず、ケット数が多い場合、

4

ケットどうしのコリジョン(衝突)によってケットが完全に消失してしまう場合がある。また符号ビット誤りが伝送上のエラー等によってある閾値、例えば50%程度に達した場合、そのケット情報は全て失われたものとし、ケット消失と判定される。さらに、ケットの到着遅延がゆらぎ吸収バッファよりも大きい場合にケット消失と判定される。これらの原因によってケットが消失し、音声の品質劣化が生じる。品質の劣化によって聴覚に不快感を与えないために、失われたケットの部分は別の何らかの信号で補償する必要がある。符号化方式によっては、過去の音声の特徴量を用いて符号化しているため、一度ケットが消失すると、復帰後しばらくは品質が劣化することがある。その品質の劣化を目立たないように補正することもケット消失補償に含まれる。例えば、復号器において、前のケット情報を用いて、パラメータの補間や音量の制御を施すことにより、たとえ一部のケット情報が欠落しても、できるだけ劣化を抑えるように処理をする。この処理は、ケットが消失したという情報が利用可能であることが条件であるが、伝送条件の悪い場合、つまりケット消失が起こりやすい場合には、消失補償による劣化抑制処理の効果は非常に大きい。

【0003】IP通信ではケットを送信しても、ネットワークの状況によって、ある程度は届かない可能性がある。IP通信では送信したケットの順番を判定し、復号器側バッファ(ゆらぎ吸収バッファ)で希望するケットが再生すべき時に到着していないと判断された場合、ケット消失と判定される。また伝送誤りによって、ケット消失と判断される場合もIP上で誤り判定機能を持たせることによってケット消失の判定をする。

【0004】ケットが消失した場合の解決策として、現在までにいくつかの手法が提案されてきた。低ビットレートの音声符号化に使用されるCELP(Code Excited Linear Prediction: 符号励振線形予測)方式でのケット消失補償では、ケット内の音声信号を周期的成分と非周期的成分に分析しておき、消失ケットに格納された信号波形のピッチ周波数が周期性であれば、適応符号帳の励振信号を用い、非周期性であれば、白色雑音をランダムに使用するという手法がよく用いられる。その他にも合成フィルタ係数を反復させる、適応・固定コードブックゲインを減衰させる、ゲイン予測を減衰させるという手法があげられる。また、PCM(Pulse Code Modulation)のような波形符号化の場合は、過去の信号からピッチ周期を解析し適当な波形を取り出し、それを繰り返すことによって、擬似的な信号を作る手法がある。この波形繰り返し補償で最も劣化の原因となりやすいのは波形の不連続によるものである。その波形の不連続が発生しやすいのは消失ケットの代わりに生成された補償信号と前後のケットの信号波形との繋ぎ合わせの部分である。この不連続性を目立たなくするために、ピッチ周期

を消失から復帰後と連続になるように調整する、あるいはOLA(Overlap add)によって、合成信号と復帰後の信号を除々に変化させていくという手法がある。また、連続でパケットが消失した場合(バースト消失)、合成信号のパワーを除々に減衰させることにより、聴覚に不快にならないような工夫をしている。

【0005】これらの手法は聴覚に不快な信号を抑制する効果に関しては有効な手法であった。しかし、あくまで擬似的な合成信号の再生であり常に原音に近い音を再生することが困難である場合が多い。パケット間において、ピッチやパワーが急速に変わったりする場合、あるいはピッチ間隔の不一致による波形の不連続性や無理な調整によって音質が著しく劣化する場合があった。

【0006】

【発明が解決しようとする課題】本発明では、従来のパケット消失補償技術の欠点を解消し、パケット消失による音声の品質劣化を改善することを課題としている。従来技術ではパケットが消失している区間で、急激な音声信号の変化によって、劣化が目立つことがあった。また常にパケット消失に備えて前後のフレームの符号を補助情報として付加すると帯域の有効活用はできない。本発明では補助情報を効率よく付加し、パケットが消失することによる音声信号の劣化を抑えることのできるパケット構成方法及び装置、パケット構成プログラム、並びにパケット分解方法及び装置、パケット分解プログラムを提供することを課題としている。

【0007】

【課題を解決するための手段】上記課題を解決するために、本発明のパケット構成方法及び装置は、音声信号をフレームごとに符号化した符号をパケットに格納するパケット構成方法及び装置において、現フレームの音声信号の繰り返しまたは該符号の特徴量の外挿により前フレーム及び後フレームの補償信号を合成し、前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より大きく、後フレームの信号波形と前記後フレームの補償信号との歪が所定の閾値より大きい場合、現フレームと、前フレームと後フレームの符号を含めてパケットを構成し、前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より大きく、後フレームの補償信号との歪が所定の閾値より小さい場合、現フレームと前フレームの符号と前フレームを示す符号を含めてパケットを構成し、前フレームの信号波形と前記前フレームの補償信号との歪が所定の閾値より小さく、後フレームの信号波形と前記後フレームの補償信号との歪が所定の閾値より大きい場合、現フレームと後フレームの符号と後フレームを示す符号を含めてパケットを構成することを特徴とする。

【0008】また、本発明のパケット分解方法及び装置は、パケット毎に格納されたフレーム毎の符号を復号化して音声信号を再生するパケット分解方法及び装置にお

いて、パケットが消失したか否かを判定し、現パケットが消失した場合、前パケットが後フレーム符号を含むとき、当該後フレーム符号を復号して現パケットの音声信号を再生し、後パケットが前フレーム符号を含むとき、当該前フレーム符号を復号して現パケットの音声信号を再生し、前パケットも後パケットも前フレーム符号も後フレーム符号も含まず現フレーム符号を含むとき、前後いずれか一方のパケットの当該現フレーム符号の復号信号の繰り返し又は該信号の特徴量の補間もしくは外挿により現パケットの音声信号を再生することを特徴とする。

【0009】

【発明の実施の形態】本発明では符号器において入力された音声信号を一定のサンプル数のフレームごとに符号化を行う。現在注目する区間を現フレーム、その信号を符号化するメインエンコーダを備え、これによって符号化されたデータをメインコードと称す。なお、現フレームよりも時間的に直前または直後のフレームをそれぞれ前フレーム、後フレームと称す。それらの信号をそれぞれ符号化する前サブコーデック、後サブコーデックを備える。符号化されたデータを各々前サブコード、後サブコードと称す。後フレームは現フレームよりも時間的に未来の信号であるので、その信号を扱うには符号器側において、1フレーム分以上の入力音声信号をバッファリングする必要がある。本発明ではその各フレーム分の符号を1パケットに詰めて送信する。

【0010】VoIPをはじめとする、パケットによる音声通信においては、ネットワークの状態によって、受信側に音声パケットが送信時刻順に届くとは限らない。ネットワークの状況によって、前のパケットとの到着時間間隔が大きくなったり小さくなったりと揺らいで到着する。この揺らぎを復号化側で吸収(解消)するために揺らぎ吸収バッファが設けられる。一般にパケットに含ませる音声フレームが短ければ短いほど、1つのパケットが消失したときの音声の劣化が小さい。ただし、1つのパケットに含ませる音声フレームが短ければ短いほどオーバーヘッドの占める割合が大きい。これは、音声のパケット化して送る場合においては、音声データ以外にもIPのヘッダ、RTP(Real-time Transport Protocol)のヘッダ等がパケット毎に付加されるためである。本発明においては現フレームを符号化するメインエンコーダは高品質符号化(64kbit/s以上)を用いる。よって1パケットに含ませる音声長として10ms程度が望ましい。音声波に周期性があるため、前後のパケットにある音声信号と相関がある場合が多い。本発明ではその特性を利用し、現フレームを符号化したメインコードに、その周辺の信号から作成した後フレームの合成信号あるいは前フレームの合成信号と、後フレームあるいは前フレーム信号を比較することによって、補助情報の必要性を判断し、前、後のサブコードをパケット消失時の対策として符号

器側において付加する。

【0011】VoIpにおいては揺らぎ吸収バッファにおける最大待機時間よりも到着が遅れて届かなかったパケットは破棄されたと判断される。(パケットが消失する原因は他にもパケット同士のコリジョン、伝送上のエラー等があげられる。)復号器側では、パケット消失と判定されない場合は揺らぎ吸収バッファに蓄積されたメインコードをメインデコーダに出力しデコードする。パケット消失の判定がされた場合は揺らぎ吸収バッファに届いている後サブコードとメインコードを組み合わせたビットストリーム、あるいは後に届く前サブコードとメインコードを組み合わせたビットストリームを使って劣化の非常に少ないパケット消失補償を行うことができる。ここでメインエンコーダは圧縮率が比較的低い、高品質の符号化方式(例えばPCM, 64kbit/s)を用い、またサブコーデックにはメインエンコーダより高圧縮、そして演算量の比較的小さい符号化方式(例えばADPCM, 32kbit/s)コーデックを選ぶ。このようにすることによって、メインコードに対して少ない情報量の付加で効率の良いパケット消失補償を行うことができる。

【0012】(符号器)本発明の符号器を図1～図7を参照して説明する。図1に符号器のブロック図を示す。入力された音声信号はフレーム形成部1において例えば音声長10ms毎にフレームが形成され、パターン分類部2に入力される。パターン分類部2においては、次の処理が行われる。

(I)無音または無声子音の判断が行われる。本発明では、現フレームの無音または無声子音の判断方法として

$$SNR = 10 \log_{10} \left\{ \frac{\sum_{i=0}^{H-1} x(i)^2}{\sum_{i=0}^{H-1} (x(i) - y(i))^2} \right\}$$

(ここで、 x : 入力信号、 y : 合成信号、 l : 1フレーム間のサンプル数を表す。)

$$CD = \frac{10}{\log 10} \sqrt{2 \sum_{i=1}^p (C_x^{(p)}(i) - C_y^{(p)}(i))^2}$$

(ここで、 $C_x^{(p)}(i)$ 、 $C_y^{(p)}(i)$ はそれぞれ入力、合成信号の p 次のLPCケプストラム i 次の係数を表す。)

【0014】無音、無声子音、波形歪の測定によって次に挙げる6つのパターンに分類できる。

- (1)無声区間である。
- (2)無声子音である。
- (3)現フレームから前フレームの波形も後フレームの波形も作ることができる。
- (4)現フレームから前フレームの波形を作ることができ

フレームにわたって波形の振幅が予め定められた閾値(例えば、量子化16ビットのうち256(2^8))以下の場合をもって判断する。なお、無音区間は周知の手段により検出することができる。

(II)無音または無声子音区間と判断されない場合(有声音区間)は、

(i)現フレーム、(あるいは前フレームと後フレーム)から“波形繰り返し補償”により前フレーム、後フレーム(現フレーム)の合成波形(信号)を生成する。合成信号列の具体例については近過去の信号のピッチ成分を抽出し、それを繰り返して外挿する。ただし、フレーム間の波形のつなぎあわせの部分は不連続とならないように重ね合わせる(OLA:Overlap add)。

(ii)それぞれの合成波形と前フレームと後フレーム(あるいは現フレーム)の波形と比べてどの程度の波形歪があるかを調べる。前フレームのコードを伝送する必要性の有無(現フレームから前フレームの波形を作ることができる)、後フレームのコードを伝送する必要性の有無(現フレームから後フレームの波形を作ることができる)を判断する。ここで伝送の必要性は前(後)フレームの信号波形と現フレームの波形を繰り返し又は外挿補間等により波形合成した合成信号列との信号雑音比(SNR)又は歪(ケプストラム距離値(Cepstrum Distance measure: CD)等)が各々所定の閾値以上(または以下)であることををもって判断する。

【0013】SNR、CDは以下のように表される。

【数1】

るが、後フレームの波形は作ることができない。

(5)現フレームから後フレームの波形を作ることができるが、前フレームの波形は作ることができない。

(6)現フレームから前フレームの波形も後フレームの波形も作ることができない。

上記パターン情報(1)～(6)は前サブエンコーダ6-1、後サブエンコーダ7-4に入力される。現フレームはメイン

エンコーダ3でメインコードが生成されパケット構成部5に出力される。前フレームは前サブエンコーダ6-1に入力されパターン情報に基づき前サブコードが生成されパケット構成部5に入力される。後フレームは後サブエンコーダ3-4に入力されパターン情報に基づき後サブコードが生成されパケット構成部5に入力される。

【0015】パケット構成部5は、図6に示すようにメインコードに前、後サブコードを付加してビットストリーム（パケット）を構成する。図2に現フレームから前フレームの擬似信号を生成する例を示す。

（図2を参照して手順を示すと、現フレームから波形繰り返しにより前フレームの合成信号を生成し、前フレームの音声信号と比較して所定の閾値以下である場合、現フレームから前フレームの擬似信号を作ることができる。）

図3に現フレームから前フレームの擬似信号を生成しない例を示す。

（図3を参照して手順を示すと、現フレームから波形繰り返しにより前フレームの合成信号を生成し、前フレームの音声信号と比較して所定の閾値以下でない場合、この場合において、前フレームの信号を圧縮して前サブコードを生成する。）

図4に現フレームから後フレームの擬似信号を生成する例を示す。図5に現フレームから後フレームの擬似信号を生成しない例を示す。この場合において後フレームの信号を圧縮して後サブコードを生成する。

【0016】パターン(1)、(2)のように無音、無声子音は、一般に周期性の無い信号であり前後のパケットに相関がなく、繰り返しによる補間を行うと音声が悪化してしまう。また、無声子音は比較的長い時間現れることが多い。しかし、無声子音はパワーが小さく前、後サブエンコーダ6-1、7-4において量子化ビットのビット数を少なくして量子化し（例えば8bitで量子化）、サブコード（無音、無声子音コード）を出力する。つまり情報量を少なくすることによって、同じ情報量で多くのフレームの重複伝送が可能となり、パケット消失に耐性を持たせることができる。このようにしてもパワーが小さいので劣化が顕著となることはない。パターン(3)の場合、フレームが例えば欠落しても前後のパケットから補間して音声劣化のほとんどない補償を行うことができる。この場合はパケットが消失しても劣化の少ない消失補償が前後の信号によって行えるため補助情報を必要としない（つまり、サブコードは付加する必要はない）。パターン(4)の場合はメインフレームから後フレームの波形を作ることができない。よってこの後フレームの消失によって音声が悪化する可能性がある。よって後フレームをサブコーデックで圧縮し、組み合わせて送信すると良い。また、パターン(5)の場合はメインフレームから前フレームの波形を作ることができない。よってこの前フレームの消失によって音声が悪化する可能性

がある。よって前フレームをサブコーデックで圧縮し、組み合わせて送信すると良い。ここでサブコーデックに圧縮コーデックを持たせる場合は通常前フレームの内部情報を引き続き用いて符号化される場合が多い。また圧縮コーデックは演算量が多くなりコーデックの負荷が大きくなるのでサブコーデックはできるだけ演算量が少ないものを選ぶと良い。（サブコーデックで圧縮してサブコードを生成する場合の例を図3、5、6に示す。）

【0017】パケット構成部5でパケット（ビットストリーム）を構成する際に現フレームのコード（メインコード）の他に伝送する必要があるコード（サブコード）を次のように判定する。(i)前フレームだけなら前方付加、(ii)後フレームだけなら後方付加、(iii)前後両方なら両方付加、(iv)必要なし（パターン(3)）の場合にはメインコードのみとする。これにより、常に1パケットに3フレーム収容するのではなく、メインコード1フレーム分のみ、前後サブコードいずれかを加えた2フレームだけのことがある。(iii)、(iv)は情報量の大き
20
さで識別できるものの(i)、(ii)は単に情報量で識別できないので互いの違いを区別するための識別情報を符号化側で付与し、復号側で何れかの状態を区別する必要がある。

【0018】即ち、従来技術では常に「サブコードを付加する」に対して本発明では「必要がある時だけサブコードを付加する」ことによって、品質は同等でも平均伝送情報量を削減することが可能となる。例えば、PCM(64kbps)をメインコーダ（デコーダ）、ADPCM(32kbps)をサブコーダ（デコーダ）として用いた場合、1パケットに収容される情報量は、(1)前サブコード付加(32kbps+64kbps=96kbps)、(2)後サブコード付加(32kbps+64kbps=96kbps)、(3)両サブコード付加(32kbps+64kbps+32kbps=128kbps)、(4)サブコード必要なし(64kbps)、(5)無音または無声子音(32kbps)となる。符号器側にて、十分な品質をとれる場合には補助情報なしとし、補償できない場合のみサブコードによる補助情報を付与するので、サブコードを常に付加する場合と比べて、サブコーデックに圧縮率は低い
30
が演算量の小さいコーデックを使用できる。

【0019】内部状態について説明する。内部状態の生成はサブコードを例えばPCM(Pulse Code Modulation)により符号化する場合には必要としない。しかし、サブコードをADPCM(Adaptive Differential PCM)、LDCELP(Low Delay Code Excited Linear Prediction)により符号化する場合においては必要となる。ここで、内部状態とは内部状態特徴量のことで、符号化に必要な特徴量を指す。例えばADPCMでは、予測フィルタ係数、適応フィルタ係数、予測係数、ステップ幅、またLDCELPでは聴覚重み付けフィルタ係数、合成フィルタ係数、予測フィルタ係数、予測係数等があげられる。サブコードが必要であると判定されればその信号をADPCMで符号化するための

内部状態が必要となる。ADPCMは量子化ステップ幅と予測係数の両方を適応的に逐次更新する手法であり、内部状態を生成するために、CELP符号化方式ほど多くの過去の信号を必要としない点で有利である。よって内部状態は、前サブコードにおいては、メインエンコーダ3とメインローカルデコーダ6-3で符号化、復号化して、この信号に基づいて内部状態生成部6-2において生成し、この信号と前フレーム信号により前サブエンコーダ6-1により前サブコードを生成することができる。後サブコードにおいては図7に示すように同様の操作を内部状態生成部7-3で時間軸において逆向きに符号化することにより内部状態を生成し、この信号と後フレーム信号により後サブエンコーダ7-4で後サブコードを作成することができる。このような構成によりADPCMによりサブコードを生成することができる。

【0020】パターン(6)の場合はメインフレームから後フレーム、前フレームのどちらの信号も波形を復元することはできない。よってこの後フレーム、前フレームの消失によって音声が悪化する可能性がある。このような場合は帯域に余裕があれば前フレーム、後フレームのどちらかの信号もサブコードとして出力させると良い。(図6参照)

上記のように分類する上で、ペイロード(ヘッダを除いた符号化列)がどのサブコードを含むか判別できない場合は識別情報(数ビット)も必要となる。例えば、10msの女性音声の場合において、再生信号と補間信号のSN値が正(0を閾値とする場合)であるとき良いフレームと判断した場合、前後いずれの packets から補間可能15%、無音区間40%、無声子音20%、前後 packets から補間不能25%程度になる。つまり、無音、無声子音を除く37.5%前後の packets から補間可能となることがわかる。SN値の閾値を変える、つまり歪の許容範囲を変えることによって帯域を制御することができる。

【0021】(復号器)図8を参照して復号器を説明する。復号器側では届いた packets を packets 分解部10において、補助情報、識別情報によりメインコード、後サブコード、前サブコード、無音、無声子音コードに分配する。メインコードはメインデコーダ11に入力され復号される。また、前、後サブコードはそれぞれ前、後サブデコーダ14-1, 15-1に入力される。

【0022】図9に示すように、

パケットロス時のサブコードが無音、無声子音コード

であれば、メインデコーダ11の符号化に用いた量子化ビット、すなわち、少ない量子化ビット(例えば8bit)に戻して再生する。

パケットの消失がない場合はメインコードを再生する。

パケットが消失した場合は、後サブコード、前サブコードの場合は符号器と同様な手法によって内部状態生成部14-2, 15-2により内部状態を生成する。メインデコーダ11でメインコードをデコードし、この信号に基づき内部状態を得ることができる。また、その信号を用いて前サブデコーダ14-1、後サブデコーダ15-1によりサブコードを復号することができる。なお、サブコードをPCMで符号化した場合には内部状態の生成は行わない。

【0023】パケットロスがあれば、出力コントローラ12で前後に対応する前フレームまたは後フレームに対応するサブコードをサブデコーダで復号して復号音声を再生する。該当信号がなければ以前の消失のない packets のメインコードに基づいて波形合成による消失補償、例えば上記のように繰り返し波形を合成して重ね合わせ再生する。つまり出力コントローラにおいて、パケットが消失した時点で、

・既に入力した packets にサブコードがある場合、該サブコードに基づく信号を再生する。

・次に入力した packets にサブコードがある場合、該サブコードに基づく信号を再生する。

・どちらの packets にもサブコードがない場合、過去の復号信号を用いた波形繰り返し補償を適用するように制御する。ここで、出力コントローラは揺らぎ吸収バッファに上記のような判別機能が含まれているものである。このような構成にすることで常にサブコードを付加する場合と比べてほぼ同等の情報量で、演算量が遙かに改善されかつ品質の良い復号音声信号を得ることができる。特に packets が2つ連続で消失した場合においても図9のように消失補償がされ品質向上が期待できる。なお、上記の例においてメインコードは packets 毎に格納されているが、1つの packets に複数フレーム分のメインコード、サブコードを格納することは任意である。

【0024】本発明手法と従来手法の符号化方式による平均ビットレートの例を示す。(メインコーデックはG.711(PCM)符号化方式(64kb/s)、サブコーデックは演算量が比較的小さいG.726(ADPCM)方式(32kb/s)を用いる。)

【表1】

Method	メインコード	後サブコード	前サブコード	平均ビットレート
本発明	PCM μ law 64kbit/s	ADPCM 32kbit/s (25%)	ADPCM 32kbit/s (25%)	80kbit/s
従来法	PCM μ law 64kbit/s	LD-CELP 16kbit/s (100%)	—	80kbit/s

(%)はそのサブコード付加率を示す。)

ITU-Tで勧告された客観評価法PESQ(Perceptual Evaluation of Speech Quality)を実施して以下の結果が得られた。ただし、従来法として特願2001-18541の発明と比較し、いずれも単一パケット消失率、2連続パケット消失率3, 5, 10パーセントのときのPESQ値を表に示す。そ

PESQ score	Random packet loss			Random 2burst packet loss		
Loss rate	3%	5%	10%	3%	5%	10%
本発明	3.996	3.873	3.624	3.904	3.588	3.234
従来法	3.849	3.763	3.618	3.622	3.384	3.058
補償無し	3.502	3.272	2.775	3.462	2.934	2.288

【0025】本発明の符号器及び復号器は、CPUやメモリ等を有するコンピュータと、アクセス主体となる端末と、記録媒体とから構成することができる。記録媒体は、CD-ROM、磁気ディスク装置、半導体メモリ等の機械読み取り可能な記録媒体であり、ここに記録されたパケット構成プログラム及びパケット復号プログラム制御用プログラムはコンピュータに読み取られ、コンピュータの動作を制御し、コンピュータ上に前述した実施の形態における各構成要素を実現する。

【0026】

【発明の効果】本発明によれば、従来法に比較して、少ない情報量の付加で、パケット消失による品質の劣化を抑え、原音に忠実な消失部分の補償をすることが可能となる。また、演算量に関しても軽くすることが可能となり、メインであるフレームの前後の補助情報をもつため、パケットが連続で消失した場合においても従来手法よりも優れた性能を発揮する。

【図面の簡単な説明】

【図1】本発明の1実施例である符号器のブロック図。

【図2】現フレームから前フレームの擬似信号を生成する例を説明するための図。

【図3】現フレームから前フレームの擬似信号が生成しない例を説明するための図。

【図4】現フレームから後フレームの擬似信号が生成する例を説明するための図。

の結果、本発明では補償無しの場合は勿論、従来法よりも高いPESQ値（PESQ値が高い方が主観的品質に優れる）、つまり主観評価値が得られた。

【表2】

【図5】現フレームから後フレームの擬似信号が生成しない例を説明するための図。

【図6】符号器のパケット構成を説明するための図。

【図7】後サブコードにおける内部状態の生成を説明するための図。

【図8】本発明の1実施例である復号器のブロック図。

【図9】復号器の機能を説明するための図。

【符号の説明】

- 1・・・フレーム形成部
- 2・・・パターン分類部
- 3・・・メインエンコーダ
- 4・・・パケット構成部
- 6・・・前サブコーデック
- 6-1・・・前サブエンコーダ、6-2・・・内部状態生成部、6-3・・・メインローカルデコーダ
- 7・・・後サブコーデック
- 7-1・・・後メインエンコーダ、7-2・・・メインローカルデコーダ、7-3・・・内部状態生成部
- 10・・・パケット分解部
- 11・・・メインデコーダ
- 12・・・出力コントローラ
- 14・・・前サブコーデック
- 14-1・・・前サブデコーダ、14-2・・・内部状態生成部
- 15・・・後サブコーデック
- 15-1・・・後サブデコーダ、15-2・・・内部状態生成部

音声信号

現フレーム

前フレーム

Z000

合成信号

①前フレームから前フレームの差分を計算して差信号を生成する

②前フレームから前フレームの差分を計算して差信号を生成する

音声信号

頭フレーム

脚フレーム

圧縮

前サブコード

合成信号

bad

②合成信号と脚フレーム信号からSNR又はCD誤差を算出して検出値と比較するといふと判断される

③前フレームが消失した場合を想定して脚フレームの波形推定により合成信号を生成

①前サブコードから前サブコードを生成

音声信号

後フレーム 現フレーム 前フレーム

③ 現フレームから後フレームの間の音声信号をつくる事ができる

④ 合成信号と後フレーム音声からS/N比はODDを計算して雑音と比較すると良いと判断される

合成信号

① 後フレームが白化した、雑音のみで現フレームの音声信号の抽出が難しくなる

② 合成信号と前フレーム音声からS/N比はODDを計算して雑音と比較すると良いと判断される

⑤ 合成信号

⑥ 後フレームが白化した、雑音のみで現フレームの音声信号の抽出が難しくなる

音声信号

後フレーム 中フレーム 前フレーム

圧縮

後サブコード

合成信号

bad

中フレームと後フレーム信号がSNR又はDOA値を計算して悪質と判断すると良いと判定される

後フレームから後サブコードを生成

後フレームが不良した信号を判定して中フレームの信号を取り戻しより合成信号を生成

图5

【図6】

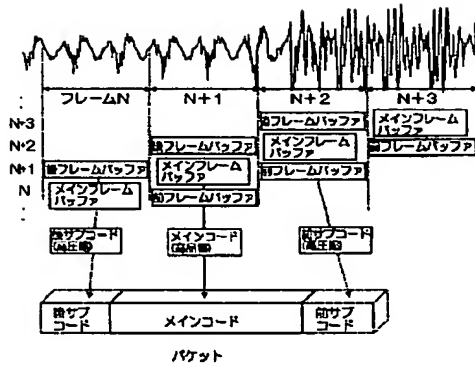


図6

【図9】

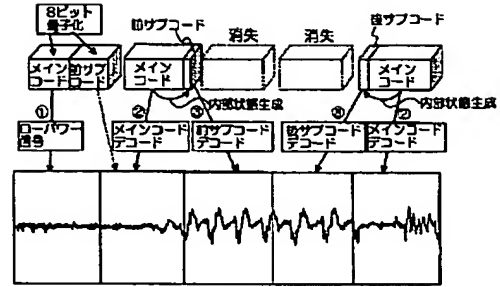


図9

【図7】

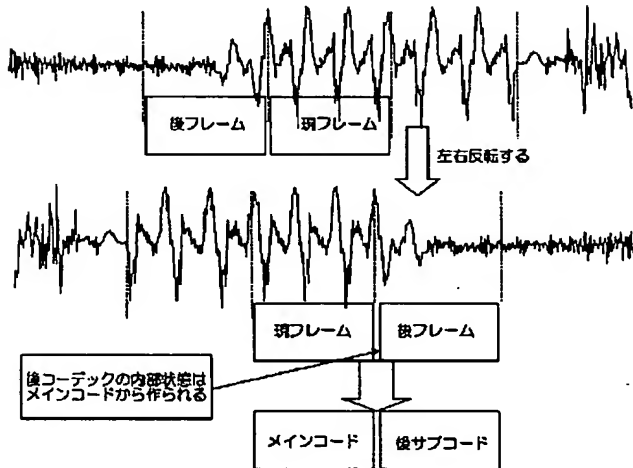


図7

【図8】

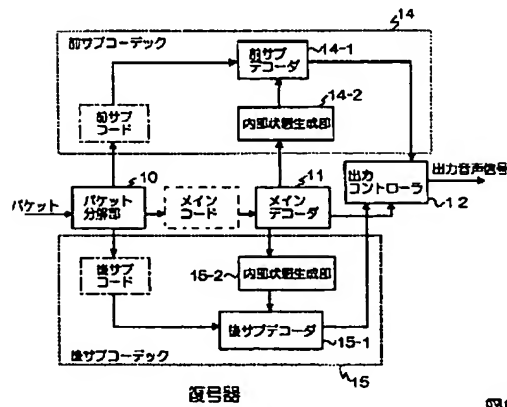


図8

フロントページの続き

Fターム(参考) 5D045 DA20

5K014 AA01 AA02 DA00 FA06 FA14

5K030 HA08 HB01 JA05 KA19 LA01

LA06 MB13